

- [Contents](#) |
- [Author index](#) |
- [Subject index](#) |
- [Search](#) |
- [Home](#)

## Factores para la adopción de *linked data* e implantación de la web semántica en bibliotecas, archivos y museos

[Tomás Saorín](#)

Universidad de Murcia, Grupo de Investigación en Tecnologías de la Información,  
Campus de Espinardo, 30071, Murcia, España

[Fernanda Peset](#) y [Antonia Ferrer-Sapena](#)

Universidad Politécnica de Valencia, Departamento de Comunicación Audiovisual,  
Documentación e Historia del Arte, Camino de Vera, s/n, 46022, Valencia, España

### Abstract

**Introducción.** Se elabora un marco para identificar el nivel de penetración y madurez de las tecnologías linked data, como una faceta específica del modelo abierto del W3C para la web semántica, en el sector de las bibliotecas, archivos y museos.

**Método.** Se aplica el marco conceptual de la teoría clásica de la difusión de innovaciones, junto con los patrones de análisis del ciclo de vida de tecnologías (Hype Cycle) desarrollados por la consultora Gartner Group. A partir de ellos compone un modelo de análisis de factores globales, adaptado a la naturaleza de la tecnología linked data.

**Resultados.** Contemplando de forma conjunta los diferentes factores de adopción de linked data, obtenemos un marco para identificar el nivel de madurez de la tecnología. Nuestro análisis indica que, pese a la intensa actividad que se detecta en el sector de bibliotecas y archivos, aún no hay un equilibrio suficiente entre publicación, consumo, regulación y

Conclusiones. Se propone un marco para la extensión de modelos linked data para aplicaciones de consumo de datos, explotar su dimensión colaborativa y tejer colaboraciones con entidades de fuera del sector LAM (*libraries, archives y museums*). El liderazgo actual de grandes instituciones debe extenderse y estimularse la innovación abierta.

[Abstract in English](#)

CHANGE FONT

## Introducción

Reflexionamos sobre el proceso de adopción de estándares para el mundo web, y cómo han madurado las recomendaciones sobre web semántica desde sus primeras formulaciones en torno a 2001-2002.

La influyente voz de Tim Berners-Lee afirmó que el movimiento de datos abiertos impactaría globalmente en el año 2010 ([Berners-Lee 2010](#)). A partir de estas fechas es el término *linked data* el que focaliza el interés conjunto del sector tecnológico y el del conocimiento abierto y libre. *linked data* es la cara tangible en el mercado de la información del concepto globalizador de web semántica.

*Linked data* está irrumpiendo con fuerza en la terminología de nuestro sector. ¿Se trata de una nueva *buzzword*, una etiqueta más para la impulsar las ventas en tecnologías? ¿Debemos prestarle una atención diferenciada del término *web semántica*? En definitiva, queremos clarificar si las tecnologías asociadas a *linked data* constituyen una verdadera transformación que afectará a la forma en la que se diseñarán los sistemas de información, la integración de fuentes de datos, la producción social de conocimiento y la organización de información digital.

Se parte de la hipótesis de que el movimiento de datos enlazados ( ) está alcanzando desde la segunda mitad del año 2011 la masa crítica de contenidos, tecnologías, visibilidad, usuarios y aplicaciones para situarse en el inicio de la fase de adopción a gran escala. En la actualidad contamos con los estándares básicos de la web semántica –rdf, owl, sparql, lod- que permitirían avanzar en el ciclo de vida de los datos enlazados: extracción, creación, enriquecimiento, enlazado, visualización y mantenimiento. Y la pregunta que nos planteamos es si acabará adoptándose en nuestro sector – *libraries, archives and museums*: - y cuáles serán los factores que determinarán la velocidad de penetración y amplitud de difusión.

Este ejercicio de prospectiva permitirá identificar los factores clave que pueden llevar al éxito o fracaso de iniciativas orientadas a *linked data* y la web semántica, ofreciendo una orientación a los responsables de estrategia tecnológica de centros, servicios de información y otros agentes del macrosector de la comunicación cultural.

Para ello delimitamos el papel de *linked data* en la web semántica, es decir, seleccionamos las iniciativas concretas que han de ser tomadas en cuenta en nuestro análisis, ya que nos interesa concentrarnos en una serie limitada de elementos, pero con gran capacidad de interconexión y efectos sinérgicos. Para su análisis, la teoría de la difusión de la innovación ([Rogers 2003](#)) nos permite contar con un marco conceptual inicial sobre el que estructurar los datos y hechos recogidos de esta observación directa de, en nuestro caso, el sector de los datos enlazados. Nos permite identificar los ciclos de vida del proceso de innovación y matizar los modelos lineales mediante la introducción de discontinuidades como las que plantean Moore y la consultora Gartner Group.

Con estos mimbres, y los instrumentos del macro-análisis estratégico de negocio para identificar los factores e interdependencias de los factores políticos, económicos, sociales, tecnológicos, ambientales y legales, elaboramos un modelo coherente de factores en tensión y equilibrio, que permita pronunciarse con mayor seguridad sobre la penetración de las tecnologías *linked data*.

## linked data: el frente más activo de la web semántica.

Dado que en este trabajo hemos puesto el foco en los *datos enlazados* realizaremos un rápido vistazo a los hitos de su formulación. Desde la primera versión de Resource Description Framework, (RDF) en 1999 hasta la liberación en 2011 del paquete LOD2 ([Verdonckt 2011](#)) para la gestión integral del ciclo de vida de *linked data* se ha experimentado un progresivo interés desde múltiples instancias. Pero sin duda ha sido clave el papel del creador de la Web, Tim Berners-Lee, como líder del W3C, organismo impulsor del desarrollo abierto de la web. Dos de las líneas centrales del W3C son los estándares abiertos y los metadatos. Frente a otras figuras de referencia, como Miller, Heath, Hendler o Bizer, hemos optado por la de Berners-Lee porque combina al mismo tiempo el liderazgo entre los tecnólogos con el de la proyección en la esfera política y mediática. Sus manifestaciones cuentan, a priori, con un mayor impacto. En el recuadro 1 recogemos algunos de sus posicionamientos específicos sobre *linked data* en el foro TED de amplia difusión.

### Recuadro 1. Tim Berners-Lee, de la web semántica a la web de los datos

En febrero de 2009 tiene lugar la conferencia pública en el influyente foro de tendencias tecnológicas TED (Technology, Education & Design: Ideas worth spreading. Berners-Lee ([2009](#)) solicita que se publiquen datos con el lema 'Raw data now!', desgranando la importancia del concepto 'Linked Data'. En septiembre de 2009 se convierte en asesor del gobierno británico, donde influirá en la política de datos públicos abiertos y enlazados.

En la TED Talk del año siguiente, en 2010, lleva la idea hacia los 'Linked Open Data', tratando de demostrar su efecto en la producción de servicios de gran potencia e impacto. Afirma que 2010 será 'the year Open Data went worldwide'.

La semilla de la web semántica existía desde las primeras formulaciones de la web, pero ésta creció sobre un modelo mucho más sencillo de enlaces indiferenciados. No obstante se estaba trabajando en el 'Semantic web Roadmap' (1998). Podemos hablar de una segunda fase a partir del impacto combinado del libro *Weaving the Web* ([Berners-Lee y Fischetti 1999](#)) y del artículo en la revista Scientific American ([Berners-Lee et al. 2001](#)) hasta la formalización a partir del lanzamiento en 2007 del Linking Open Data community project en el W3C, dentro del grupo de interés SWEO (Semantic Web Education and Outreach), cuando se acuña el término definitivamente. Observamos cómo aparece *linked data* como una parte especializada de la web semántica, enfocada a la publicación de datos estructurados en RDF usando Uniform Resource Identifier (URI), dejando a un lado, temporalmente, las ontologías e inferencias. Esta simplificación reduce las barreras de entrada a los proveedores de datos, facilitando la extensión de su adopción ([Hausenblas 2009, 5](#)). Al igual que ha ocurrido en casos

anteriores, como TCP/IP o RSS, una cierta simplificación de requisitos genera un gran impacto al generalizarse su rápida difusión.

Podemos ver que desde la formulación genérica de web semántica, como diagnóstico y tratamiento global a los problemas estructurales de la web, se camina hacia una formulación práctica: los datos enlazados. Esta fórmula para generar las condiciones para la web semántica a través de los datos distribuidos da lugar al término 'web de los datos'. También O'Reilly (2005) hace de los datos en la web uno de los elementos constitutivos de la Web2.0: *'los datos son el nuevo Intel inside'*. Heath y Bizer (2011) hablan del *'data deluge'* en el que si antes *'la World Wide Web ha revolucionado la forma en que conectamos y consumimos documentos, ahora está revolucionando la forma en que descubrimos, accedemos e integramos y usamos datos'*.

La web semántica aparentemente va materializándose siempre de forma incompleta y distribuida, y a veces contradictoria. Esta es la forma natural de la web;

utilizar términos como «objetivos», «planificación» o «metas» en el ámbito de la web es un sinsentido, ya que las necesidades o ideas nuevas, precisan de herramientas que se crean y perfeccionan sobre la marcha, fruto de la obtención de resultados tras su aplicación (Pastor 2011, 19).

Establecer una analogía con el movimiento Open Access permite apreciarlo con claridad. Si bien los repositorios basados en el protocolo Open Archives Initiative (OAI) son la parte más visible y homogénea, se fueron desarrollando desde acciones de muy diversa índole, que están conduciendo a un éxito incontestable. Y como señala Suber (Poynder 2011) el proceso ha sido orgánico y no planificado: no ha existido una estructura formal ni jerárquica, ni un líder indiscutible, imponiéndose al lobby de las editoriales establecidas.

## Modelos para analizar el éxito y la difusión de las innovaciones tecnológicas

### Los efectos de red

Castells (1996) señala el enfoque en red en el capitalismo informacional como una forma específica para analizar grupos y fenómenos sociales. Las redes son estructuras abiertas, capaces de expandirse sin límites mientras puedan comunicarse entre sí. Por lo tanto una estructura social basada en las redes es un sistema muy dinámico y abierto, susceptible de innovarse sin amenazar su equilibrio. En las redes intensivas en comunicación multidireccional, el fenómeno denominado capilaridad estimula la innovación (Cornellá y Flores 2007).

Las formas de crecimiento de la propia web, y de muchas de sus aplicaciones, han traído a primer plano el llamado *network effect*, que se podría explicar como el incentivo creciente por participar en un fenómeno que se va extendiendo asimétricamente en una red. *'The more people who are playing now, the more attractive it is for new people to start playing'*. Allemang y Hendler (2008, 8-9) identifican este modelo como clave para la expansión de la web semántica. En su estudio de las leyes económicas de la información, Shapiro y Varian (2000) identifican que el feedback positivo en las tecnologías sometidas a fuertes efectos de red suele tener una primera fase relativamente lenta de asentamiento, seguida de un crecimiento explosivo.

En economía se denominan 'externalidades de red' a los efectos que hacen que el valor de un producto o servicio para un usuario dependa no sólo del producto o servicio en sí mismo, sino del número de usuarios que lo utilicen (López Sánchez y Arroyo Barrigüete 2006: 22). Las externalidades de red se consideran un fallo de mercado, porque la cantidad de usuarios previos de un servicio limita la entrada posterior de competencia. Existen distintos tipos de externalidades de red, pero nos interesan especialmente las llamadas *'de aprendizaje'*, pues se encuentran muy relacionadas con Linked Data. Las externalidades de aprendizaje ponen el foco de atención en los conocimientos específicos que pueden tener usuarios y proveedores de servicios sobre una tecnología.

Frecuentemente se menciona la Ley de Metcalfé que cuantifica el efecto de red para espacios interconectados, como los fenómenos web 2.0 y Web semántica (Hendler, Golbeck 2008, 14). Metcalfé enuncia que mientras el coste de una red crece linealmente con el aumento del número de conexiones, su valor es proporcional al cuadrado del número de usuarios. Aunque los términos matemáticos de la ley han sido sometidos a controversia, la idea central del valor de la interconexión mantiene una gran fuerza explicativa en diferentes contextos de uso. Se señala que la web de datos no puede perder de vista el adjetivo *'enlazados'*, que es el que aporta sustancialmente su valor. La misma dependencia de la capacidad de ser enlazados explica los ciclos de consumo y producción de los contenidos generados por usuarios, especialmente *'where the Social Web meets the Semantic Web'* (Hendler and Golbeck 2008: 17).

### Modelos descriptivos y prospectivos de adopción de tecnologías

El conocimiento de los procesos de innovación es un área de interés para las políticas científicas e industriales ya que en mercados sometidos a la competencia global, el liderazgo tecnológico constituye un valor económico. No hemos de ver la innovación simplemente como una mejora tecnológica, sino como la explotación con éxito de nuevas ideas de forma sostenible, incluyendo casi tanto más que la invención, la comercialización o implementación.

En el anterior apartado hemos destacado el liderazgo y la visibilidad de Berners-Lee, puesto que las expectativas son una parte de los atributos de la innovación. Otros atributos que influyen en la adopción de una nueva tecnología son, conforme al modelo de Rogers de difusión de innovaciones (2003):

- Ventaja: grado en que una innovación es percibida como buena idea.
- Complejidad: percepción de la dificultad de entendimiento de uso.
- Compatibilidad: capacidad de pervivir con los valores existentes y el sistema social.
- Experimentación: capacidad de formar parte de un plan y ser probada.
- Visibilidad: grado en que los resultados son visibles a otros.

Del modelo de Rogers suele usarse su parte más visible: la representación gráfica mediante una curva que combina el factor temporal junto con la cantidad de usuarios que adoptan una tecnología, estableciéndose varias tipologías de usuarios. El número de usuarios en un momento dado permite describir el estado de implantación de una tecnología y analizar las evoluciones en el ciclo de vida. En diferentes presentaciones y documentos sobre la web semántica se han usado estas curvas y los tipos de usuarios asociados a sus fases: Innovadores, Primeros adoptantes, Mayoría precoz y Mayoría rezagada. La figura 1 muestra la percepción del W3C en la que se observa que las tecnologías semánticas aparecen aún en la fase ascendente de un mercado precoz.

Sobre estas curvas hay autores que introducen también las discontinuidades que originan cambios de estado, con el fin de aportar una visión más intermitente de la evolución en la implantación de innovaciones tecnológicas (Moore 2008). Moore señala un punto crítico en la transición hacia la "Mayoría precoz", que denomina "El abismo" (The chasm), donde cada factor puede ser crítico para la continuidad del crecimiento o su estancamiento.

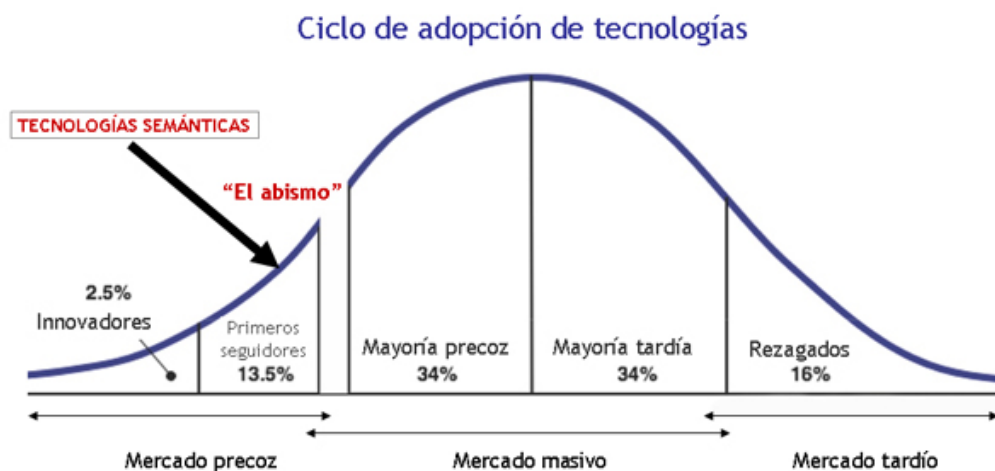


Figura 1: Ubicación de la web semántica en el ciclo de vida de la adopción de tecnologías. Fuente: Adaptación de figura de Ivan Herman (2011)

La importancia de estas discontinuidades fundamenta otros modelos, como el de la prestigiosa consultora tecnológica Gartner Group, que introduce en la curva el concepto de expectativas o hype cycles (Fenn y Raskino 2008). La metodología de análisis hype cycles ofrece una representación gráfica de la madurez en la adopción de tecnologías y aplicaciones, así como su capacidad para ser potencialmente relevantes en la resolución de problemas reales de las empresas y en explotar nuevas oportunidades. En lugar de representar un crecimiento continuado, las curvas muestran un pico de crecimiento muy acelerado, conocido como *sobreexpectación*, que corresponde a una fase inicial de euforia tecnológica.

Las fases delimitadas en este modelo son:

- *Lanzamiento* (Technology trigger): presentación del producto o cualquier otro evento, por lo que genera interés y presencia en los medios.
- *Pico de expectativas sobredimensionadas* (Peak of inflated expectations): el impacto en los medios genera normalmente un entusiasmo y expectativas poco realistas. Es posible que algunas experiencias pioneras se lleven a cabo con éxito, pero habitualmente hay más fracasos.
- *Abismo de desilusión* (Trough of disillusionment): el interés por la tecnología decae con fuerza porque no se cumplen las expectativas. Por lo general la prensa abandona el tema.
- *Rampa de consolidación* (Slope of enlightenment): aunque el foco de atención se haya desplazado a otras tecnologías, algunas empresas continúan su esfuerzo por entender los beneficios que puede proporcionar.
- *Meseta de productividad* (Plateau of productivity): los beneficios se han demostrado y aceptado. La tecnología se vuelve cada vez

más estable y evoluciona en segunda y tercera generación. La altura final de la meseta varía en función de si la tecnología es ampliamente aplicable, o de si sólo beneficia a un nicho de mercado.

Su perspectiva evolutiva la hace útil para la planificación adaptada a contextos prospectivos:

- Separar las *estrellas rutilantes* (hype) de aquellas que ofrecen una promesa de tecnología en fase comercial.
- Reducir el riesgo de las decisiones de inversión en tecnologías.
- Comparar la comprensión Internet del valor de la tecnología para una compañía con análisis objetivos realizados por expertos en tecnologías de la información, que pueden complementar el punto de vista limitado a un sector o a una problemática.

Este modelo de representación también permite contemplar si los adoptantes de una tecnología se han precipitado, han abandonado demasiado pronto, la han adoptado demasiado tarde o, incluso, si llevan demasiado tiempo descolgados de una tecnología madura y con buenos resultados demostrados.

Sus informes anuales cubren más de 75 áreas temáticas y 1800 tecnologías. Nos interesa prestar atención al informe monográfico 'hype cycle for Web and user interaction technologies' de 2010 y 2011.

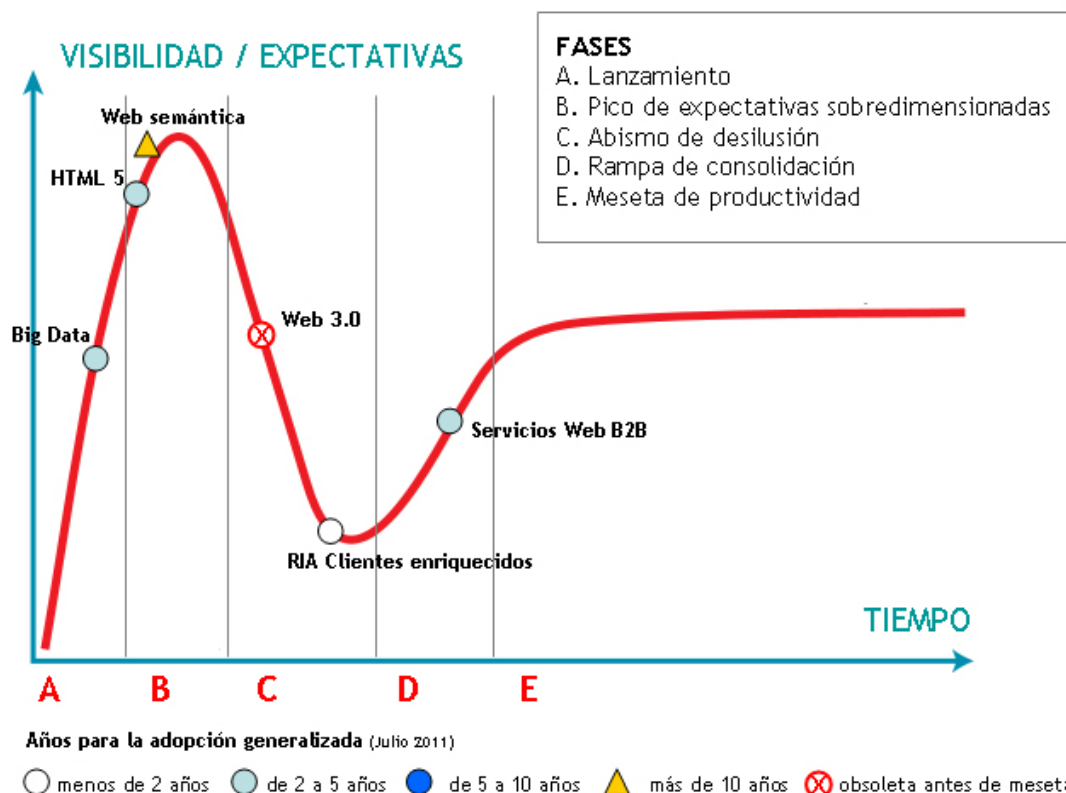


Figura 2. Adaptación y selección de elementos del 'Hype cycle for Web and user interaction technologies'. Gartner Group, 2011'

En ellos sigue identificándose 'web semántica', mientras que *linked data* aún no ha entrado como etiqueta ni en éste ni en los otros informes de la serie hype cycle (por ejemplo *master data management*, *data management* o *enterprise information management*). La Web semántica se recoge específicamente en sus informes desde 2008 (el año anterior se denominaba *Web services semantic standards*), y se situaba en fase de caída. Sin embargo a partir del año siguiente, 2009, vuelve a tomar impulso y en lugar de seguir el proceso completo, vuelve a situarse iniciando el ascenso hacia el pico de sobreexpectación. Esta sería su cronología:

	Hype cycle 2010	Hype cycle 2011	Plazo para implantación a gran escala
<b>Web semántica</b>	Ascendiendo la última parte de la curva de sobreexptación	En el máximo pico de sobreexpectación.	Más de 10 años

Gartner considera en 2011 estas tecnologías aún como adolescentes, con una penetración en el mercado entre el 1-5 %, pero con altas posibilidades de beneficio a largo plazo, donde pueden desarrollarse modelos de negocio sostenibles, en especial en los campos de la salud, ciencias bio-sanitarias, bibliotecas, defensa, servicios públicos y servicios financieros.



Las conclusiones de Gartner aconsejan valorar con cautela el impacto real de las continuas noticias sobre software para publicar-consumir RDF, proyectos que usan la etiqueta linked data y conjuntos de datos en RDF de diferentes sectores. Los movimientos de las grandes empresas tecnológicas (ORACLE, Google, Yahoo o Microsoft, entre otras) que están apostando en muchos de sus productos y servicios por las tecnologías de la Web Semántica, e incluso desarrollando proyectos conjuntos, podrían ser vistos como una enmienda al análisis de Gartner, e indicar que la fase de declive se está dejando atrás y comenzando la madurez para un mercado. Pese a tratarse de tecnologías de propósito general, el ritmo de penetración en cada sector es diferente. En áreas como el seguimiento de grandes flujos de información en redes sociales, posicionamiento web y manejo de datos corporativos se pueden estar consolidando mercados, mientras que el sector cultural nos encontremos en una fase menos avanzada.

## Modelo para el análisis estratégico de la implantación de los datos enlazados.

El análisis estratégico se aplica a organizaciones e instituciones, con el fin de orientar adecuadamente sus objetivos a largo plazo, adaptándose a los elementos de cambio decisivos del entorno. Una comprensión rica de los factores del internos y externos del mercado permite construir con antelación suficientes estrategias defensivas, ofensivas, de supervivencia o de reorientación ([Arjonilla Domínguez; Medina Garrido 2009, 109](#)).

Para el estudio de la implantación social de las tecnologías de la información, se utilizan diferentes aproximaciones que, más allá de los datos factuales (volumen de ventas, número de empresas, hogares conectados, etc.), intentan explicar la interdependencia entre los siguientes elementos: prácticas sociales, mercado e infraestructuras.

### Factores de despegue de las tecnologías linked data

Como hemos presentado con anterioridad, la idea de web semántica es un proyecto que tiene ya un largo recorrido, más teórico y de laboratorio, que práctico y de aplicación industrial. Pedraza-Jiménez, *et al.* (2007) todavía no aparece el concepto *linked data* dentro de un estado de la cuestión sobre la web semántica, por lo que reiteramos nuestra hipótesis de partida. Desde la esfera de la web semántica, surgen las tecnologías *linked data* pasando en poco tiempo de ser consideradas un subproducto tecnológico (*spin-off*) a ser visibles autónomamente, concentrándose en ellas muchos esfuerzos, desde diferentes áreas, actuando como nodo impulsor del resto de factores necesarios para la difusión de una innovación, generando una retroalimentación positiva con efecto red.

La sociología de la tecnología nos ha enseñado a dejar de ver la evolución de la tecnología como un proceso natural de progresiva implantación. Las formas que se adoptan son resultado de tensiones entre intereses contradictorios influyendo distintos factores como son la inversión de recursos económicos, las habilidades prácticas o las formas organizativas.

En los análisis estratégicos para contextos más amplios que el de la organización individual, el enfoque PESTEL para macro-entornos ([Gillespie 2007](#)), pueden ser de utilidad, al contemplar los siguientes factores: políticos, económicos, sociales, tecnológicos, ambientales y factores legales. La utilidad de este enfoque no será tan solo la enumeración de factores, sino la comprensión de la medida en que afectan al caso de análisis, así como las tensiones entre ellos. En especial, la identificación de los elementos que actúan como facilitadores o impulsores (o al contrario, como barreras) permitirá priorizar la acción y distribuir correctamente los esfuerzos. Dado el carácter sistémico de las tecnologías linked data, consideramos conveniente adoptar este enfoque, puesto que ninguna iniciativa puede desarrollarse aislada del contexto de la red y del resto de actores que actúan en la cadena de generación y reutilización de información. Para salir del laboratorio al mercado, y ser un agente de transformación con impacto, ha de construirse una densa trama de interrelaciones a muchos niveles diferentes.

También el punto de vista de los mercados, como modelos de comportamiento económico y social en reequilibrio constante, nos ofrece ejemplos de modelos para comprender el ciclo de los productos a partir de factores interrelacionados, como la 'Brújula de posicionamiento competitivo' (*competitive-positioning compass*) de Moore (2008, 100-116) que se articula a partir de cuatro puntos cardinales: tecnologías, producto, empresa y mercado. En las fases de mercado precoz, en el que las decisiones están dominadas por un reducido número de entusiastas e innovadores, los valores clave son las tecnologías y los productos. Cuando se penetra en un mercado masivo, en el que las decisiones están dominadas por el grupo de usuarios pragmáticos y conservadores, los aspectos dominantes son el mercado y las empresas. En este contexto, la barrera que hay que cruzar, es la de la transición desde los valores centrados en el producto a los centrados en el mercado.

Existen otros modelos aplicables a la planificación de la implantación de linked data, como la Metodología RISP (Reutilización de la Información del Sector Público) del CTIC que identifica en el contexto del *open government* las fases de: Sensibilización, análisis inicial de los datos, enriquecimiento semántico, y, por último, exposición y reutilización.

Pero aunque permiten conceptualizar el proceso interno de una institución para liberar sus datos, no recogen los factores existentes en un contexto social, en el ecosistema digital. Así hace Abella (2011) al contemplar en los factores avanzados de una estrategia de reutilización el 'ecosistema', definido como la forma en que la 'entidad publicadora promueve que haya interacción entre los distintos agentes industriales que

Haremos a continuación una primera enumeración de algunos de los elementos de muy diversa naturaleza (económica, institucional, tecnológica, de servicio, liderazgo, etc.), que creemos que demuestran que estamos avanzando hacia la maduración de Linked Data, y por tanto de las tecnologías de la web semántica. Dado que Linked Data es una tecnología de propósito general, junto a los casos del sector de las bibliotecas-archivos-museos, se incluyen otros de carácter intersectorial.

### Factores económicos y de mercado

- *Reducción de costes de proceso.* Cuando una tecnología permite una sustancial reducción de los costes de producción termina imponiéndose. En el ámbito de la información, la reducción de costes puede aparecer de múltiples formas: una simplificación de los procesos de transformación, una sustitución de tecnologías complejas, o una reducción del tiempo de elaboración. Para la integración de datos, uno de los caballos de batalla de los sistemas de información, la radical normalización sintáctica y semántica que plantea RDF implica que el uso masivo de datos sea accesible a un mayor número de empresas y organizaciones.
- *Oportunidad de mercado para empresas tecnológicas.* Son necesarios modelos de negocio asociados a una tecnología que proporcione servicios de calidad. Es vital la existencia de empresas de referencia, capaces de liderar la oferta de soluciones técnicas, plataformas e innovación sostenible. Ha de existir un sector empresarial especializado que dé respuesta a una necesidad social, facilitando la contratación de servicios y la rentabilidad del cambio. Podríamos pensar en Talis o, en nuestro sector, Digibis.

### Factores tecnológicos y operativos

- *Reglas sencillas.* Ya en ocasiones se ha demostrado que los avances de la red tienden a imponerse con versiones sencillas de estándares tecnológicos. La sencillez y gradualidad de implantación forman parte del código de los datos enlazados, adoptando diferentes niveles de calidad, que permiten transiciones progresivas hacia un espacio de datos global.
- *División del trabajo.* Es uno de los principios de la economía de mercado, y en el caso de *linked data* más aún: cada agente del proceso puede ocuparse de una pequeña parcela (vocabularios, publicación de datasets, consumo), lo que facilita un crecimiento no planificado, independiente. También permite que los costes se repartan y se distribuyen en el tiempo y entre los agentes. Esa diferencia de ritmos, que a veces puede ser exasperante, sin embargo permite que funcionen las dinámicas del 'network effect'. La principal división se detecta entre productores de datos y creadores de aplicaciones (integración, presentación y utilidad-valor), es decir los consumidores de primer nivel de esos datos, que ofrecen servicios al usuario final.

### Factores de alcance social y sectorial

- *Utilidad corporativa.* A diferencia de tecnologías como OAI/PMH, rdf tiene una aplicación directa en la integración de los propios sistemas internos de las organizaciones (middleware linked data), y por lo tanto pueden existir modelos de servicio antes de que se expanda en la web pública. También en el marco B2B (Business to Business) los beneficios de Linked Data son muy tangibles, puesto que sitios web de empresas pueden integrar a bajo coste datos de otras empresas con las que medien alianzas o acuerdos de colaboración. El uso de *linked data* en ámbitos de Open Government a veces eclipsa el campo del 'Enterprise linked data' ([Wood 2010](#)), en el área de gestión de los datos, para integración e interpretación.
- *Complementariedad con otros sectores.* Cuando una tecnología apoya más de una función se multiplican los efectos de red. *linked data* puede influir en buscadores, posicionamiento, accesibilidad, redes sociales, integración de datos, portales distribuidos, etc. Además su aplicabilidad multisectorial es una de sus bazas, sobre todo para bibliotecas, que pueden ver como su mercado de proveedores de servicios se diversifica y se abarata debido a la competencia.
- *'Killer application'.* La existencia de aplicaciones conocidas y con una excelente experiencia de usuario, desvela para otras empresa el valor oculto de una tecnología, influyendo en hacerla masiva. Podríamos pensar en la adopción de rdf por Facebook a través de Opengraph como un paso en esta dirección. Aún no detectamos, en el sector de las bibliotecas, una aplicación de gran éxito, dado que las principales iniciativas aún son de publicación de datos y vocabularios. El estímulo de proyectos piloto con capacidad de demostración, como los derivados del [ProgramaDiscovery de JISC y RLUK](#) para reforzar un 'ecosistema para los metadatos', son importantes, pero de impacto aún limitado.
- *Innovación y creatividad.* La innovación se ve estimulada cuando existe un margen para experimentar de forma creativa, para diferenciarse y para construir una experiencia nueva. Un ejemplo claro son los Hackathons de *linked data* o el Desafío Abredatos ([Sierra 2011](#)). Los voluntarios que proponen sistemas de representación infográfica o geoposicionada de datos complejos permiten incorporar a los proyectos otros perfiles ajenos a las tecnologías: diseño, arte y comunicación. La creatividad tiene una dimensión aplicada, la innovación, que en nuestro marco de análisis, se presenta como innovación abierta, aportando un elemento de estímulo determinante para su difusión ([Saorín 2011](#)).

### Factores normativos y de reconocimiento

- *Exigencia reguladora.* Cuando tanto las propias tecnologías como las políticas de información van unidas a una exigencia legal en una jurisdicción (nacional, europea, federal) se facilitan los procesos de decisión. Junto a los motivos económicos y operativos para el cambio se suman los requisitos a cumplir ante los reguladores (compliance o conformidad). Lo hemos visto ya en el desarrollo de las normativas para el acceso a datos personales, en los mandatos de Open Access, en la ley de reutilización de datos y en los mandatos Government Open Data.
- *Reconocimiento y estándares de calidad.* El reconocimiento público de un esfuerzo de calidad interno proporciona un elemento de refuerzo a las decisiones que se adopten. Este refuerzo externo, junto con la superación de auditorías, permite a las organizaciones salir de sus dinámicas internas. Tanto las presiones del entorno regulador como la perspectiva de mejorar su percepción por los clientes particulares y corporativos suponen un estímulo para que las organizaciones adopten las innovaciones. La certificación pública de calidad, o de responsabilidad social corporativa, es un lubricante para el cambio organizativo, como sucede con los sistemas de calidad o la accesibilidad en la web. Alcanzar las cinco estrellas de linked data, o entrar en alguno de los mapas-directorios de iniciativas open data, da a las organizaciones una meta y un reconocimiento en los medios.

Podríamos continuar esta enumeración con otros aspectos tanto o más importantes que los reseñados hasta ahora: masa crítica de datos, cambios legislativos, reducción de barreras legales y barreras de entrada al negocio, tecnologías accesibles y baratas, inclusión en programas formativos universitarios, desarrollo de estándares, estándares de facto, diferenciación con otras tecnologías, asentamiento de encuentros técnicos monográficos, inserción en la financiación pública sobre ciencia y sociedad digital, coherencia de iniciativas descentralizadas, credibilidad y confianza, métricas, etc.

Finalmente nos gustaría volver a insistir en la importancia del liderazgo. El pequeño grupo de usuarios innovadores y adoptantes precoces de una tecnología cumplen un papel vital en la extensión de su uso, y sus posicionamientos son críticos en el paso a un mercado maduro. Para el Liderazgo personal e institucional ya hemos destacado la importancia de que *semantic web* y *linked data* estén asociadas a Tim Berners-Lee -premio príncipe de Asturias, doctor honoris causa y figura reconocible en el nacimiento de la web- lo que estimula la atención social al concepto y a las tecnologías asociadas. Habría que identificar qué figuras (e instituciones) en cada comunidad sectorial y geográfica actúan de líderes de opinión, influyendo en la agenda pública. W3C es un líder promotor del cambio, que además ha considerado a las bibliotecas en una posición muy necesaria, a través de uno de sus Grupos Incubadora sectoriales (LLD XG) [W3C 2011](#). La British Library y la Biblioteca Nacional de España están actuando con decisión en el campo de las bibliotecas, experimentando la producción de datos a partir de registros MARC. Tampoco hemos de olvidar la puesta en marcha de Centros de Investigación y redes de excelencia: ejemplo de ello son el [CTIC](#) y la [Red Temática Española](#) de *linked data* en España, el [Digital Enterprise Research Institute](#) (DERI) en Irlanda o el [Planet Data de la Unión Europea](#). A un nivel más amplio, el Data Curation Center (DCC) británico entronca al mismo tiempo con las iniciativas de datos científicos abiertos.

## Cuadrante de factores clave de proyectos Linked Open Data Bibliotecarios

La anterior enumeración de factores adquiere una forma más inteligible si tratamos de estructurarlos alrededor de los siguientes seis conceptos vinculados a los proyectos *linked open data*:

- *Definiciones de datos y vocabularios de valores.* Ontologías, conjuntos de elementos de metadatos, perfiles, registros de metadatos. En definitiva, conjuntos de datos y metadatos disponibles.
- *Conjuntos de datos (datasets).* Fuentes de datos que serán el contenido susceptible de interesar a los usuarios y de aportar valor por agregación, combinación o reutilización.
- *Aplicaciones de consumo:* servicios que combinan, aplican, reutilizan y ofrecen valor a partir de los datos.
- *Tecnologías:* plataformas que hacen sencillo publicar, consultar y enlazar información en contextos de gestión de contenidos de propósito general (CMS) y corporativos (ECM), comunicación personal, redes sociales, especializadas y corporativas.
- *Regulación:* normativa, legislación y políticas
- *Prácticas sociales y culturales.* Extensión del uso de aplicaciones que usan datos enlazados, percepción profesional y social sobre los datos abiertos y enlazados y, en sentido amplio, implantación de la cultura open data.

Estos elementos nos permiten identificar pares de complementariedades básicas presentes en los proyectos *linked open data* (Figura 3).



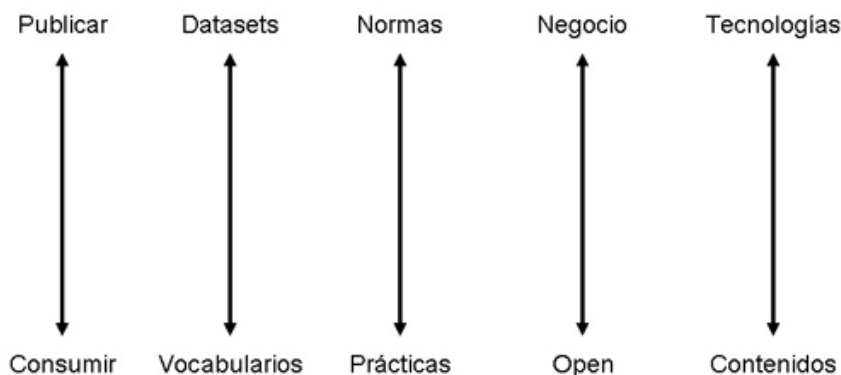


Figura 3: Pares de complementariedades en proyectos *linked open data*. Fuente: Elaboración propia

- *Publicar y Consumir*. Una fuente de datos potencialmente valiosa nunca será percibida como tal hasta que algunas aplicaciones de éxito usen sus datos para ofrecer una propuesta de valor tangible, aceptada ampliamente por los usuarios. Podría denominarse también Datasets y Aplicaciones de consumo, para hacer referencia a las tecnologías centradas en publicar datos enlazados frente a las de consulta e integración de datos de diferente procedencia.
- *Datasets y Vocabularios*. Para muchas necesidades de publicar datos enlazados pueden bastar las ontologías existentes, o usarse una combinación de elementos de varias de ellas. Se considera conveniente el máximo de reutilización de los vocabularios más extendidos, para así favorecer la conexión de datos y la facilidad de uso. Las instituciones tienen que distribuir sus esfuerzos entre publicar volúmenes significativos de información valiosa, y la definición de modelos de datos ajustados a sus necesidades. A partir de cierto punto, poblar de datos es más relevante que la creación de nuevas ontologías, y se recomienda usar ontologías ligeras.
- *Normas y prácticas*. Muchos de los avances en la web funcionan por acuerdos prácticos de mínimos, flexibles, basados en una buena relación esfuerzo-beneficio. El apoyo a ciertas normas de facto que han demostrado su funcionamiento, suele preceder a la consolidación de una norma o la regulación de un marco normativo riguroso.
- *Negocio y Apertura*. La creación de un mercado para los datos enlazados puede requerir la creación de una amplia base de uso a partir de datos abiertos de uso gratuito y libre. Las decisiones sobre el momento y el lugar para introducir costes y relaciones comerciales es delicada para el equilibrio y crecimiento del ecosistema. Muchos de los productos estrella de la web no surgieron con un modelo de negocio bien definido. Lo fueron desarrollando a partir de los patrones de uso durante su crecimiento, a través de ingresos directos o indirectos. La adopción de servicios y fuentes de datos enlazados se basa en la eliminación de las barreras económicas.
- *Tecnologías y Contenidos*. El entusiasmo de parte del sector de la ingeniería de la información por las bondades de las tecnologías de publicación y consumo de datos enlazados, tiene que ser respaldado por unos contenidos valiosos que permitan crear utilidad en servicios concretos. Las tecnologías están al servicio de datos y metadatos. Será necesario invertir recursos para liberar nuevas fuentes de datos que no están fácilmente disponibles.

Cada uno de estos factores puede tener ciclos de desarrollo y atención propios. Los desequilibrios y sinergias que se producen entre ellos nos permiten entenderlos como un ecosistema. En el sector británico de bibliotecas, archivos y museos está en marcha Discovery, a metadata ecology for UK Education & Research, basada en una combinación de licencias abiertas para estimular la innovación en servicios ([Discovery... 2012](#)). De hecho se habla de 'espacio global de datos' ([Heath, Bizer 2011](#)), y como tal espacio se entiende una construcción social dinámica multidimensional. Rufus Pollock ([2011](#)), desde la Open Knowledge Foundation usa también la conceptualización de 'Open Data Ecosystem', de forma parecida a la ecología de metadatos anteriormente mencionada. Dado que en la extensión de *linked data* no se aprecian apoyos institucionales decididos (centros que se vean obligados a implementarlos), la motivación en el mundo open se apoya en circunstancias más difíciles de identificar. Aquellos proyectos en los que coexistan factores a su favor, e interrelacionados, en todas las áreas, contarán con muchas más posibilidades de éxito y continuidad. Pretendemos por tanto ayudar a visualizar la globalidad de las acciones y actores necesarios en la adopción de los datos enlazados.

Para construir un modelo que sirva para describir el nivel de madurez de *linked data*, en un sector dado, es necesario reducir los elementos representados, de cara a conseguir mayor claridad visual. Para ello usamos la forma convencional de un doble eje, en el que se modulan 4 características, con las que se delimitan cuadrantes. Para ello se agrupa en una sola categoría los Conjuntos de datos y los Vocabularios o Esquemas de Metadatos, y en otra la Regulación y Prácticas sociales-culturales. El resultado es:

1. *Tecnologías*. Este eje refleja los aspectos estrictamente técnicos para los datos enlazados. La facilidad de acceso a herramientas de diferente tipo y alcance multiplica el alcance y la entrada de usuarios hasta alcanzar un mercado maduro.
2. *Conjuntos de datos (Datasets) y Definiciones de datos y vocabularios de valores*. La materia prima básica del ecosistema de datos enlazados es la puesta a disposición de conjuntos de datos relevantes, fiables y que usen ontologías suficientemente expresivas.
3. *Aplicaciones de consumo*. Este eje contempla los aspectos de mercado y servicios, dirigidos a comunidades de usuarios concretas, estimulando el crecimiento y mejora de los anteriores ejes.

4. *Regulación y prácticas sociales-culturales*. Este eje recoge la dimensión social que toda innovación tecnológica conlleva, tanto en los aspectos de institucionalización como en los de cambio cultural y educativo. Este eje impone limitaciones o estímulos a los anteriores, mediante normas, requisitos, conductas y estrategias.

El equilibrio entre estos 4 polos delimita una 'zona de madurez' de la tecnología *linked data*. Cada uno de los factores tiende a ser dominante en cada nuevo proyecto: identificando sus carencias en el resto de áreas podremos mejorar la gestión de un proyecto de *linked data* (por ejemplo un prototipo de aplicación que no dispone de licencia para explotar los contenidos, o la orientación de las tecnologías hacia la publicación de datos y no hacia el consumo).

Podemos tomar estas líneas de fuerzas para elaborar un mapa dividido en cuadrantes, en el que situar cada uno de los avances que se van produciendo en *linked data*. Variando su localización en el diagrama podemos tanto describir la evolución temporal de un elemento, o bien representar la transformación que hemos planificado en un proyecto de mejora.

En la Figura 4 se representa el modelo para representar la madurez de las iniciativas de datos enlazados, aplicable al sector libraries, archives and museums. Se trazan los siguientes cuadrantes:

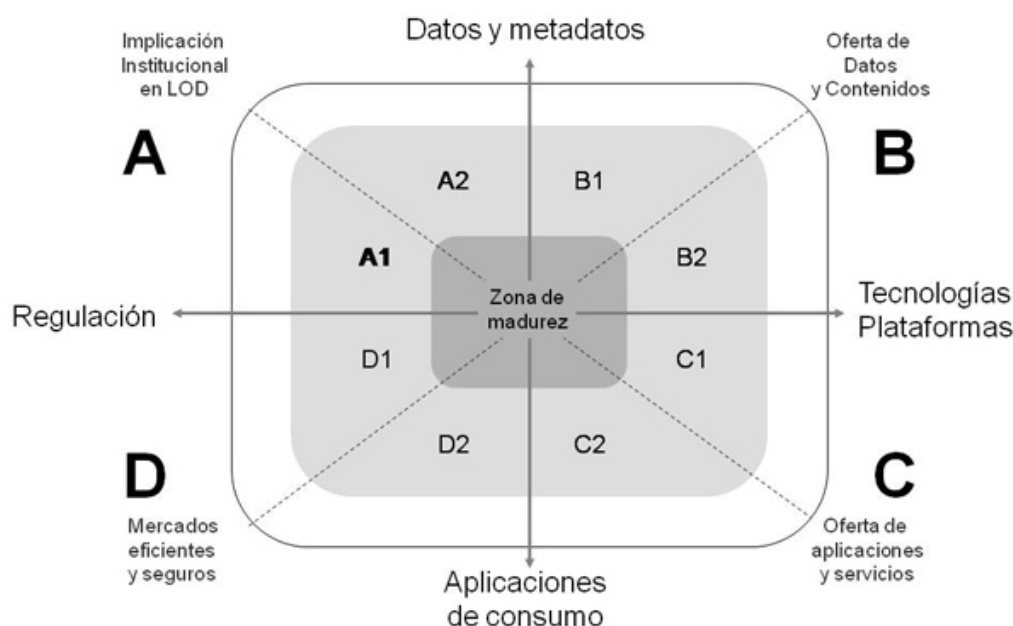


Figura 4: Cuadrantes para la madurez de *linked data*. Fuente: elaboración propia

A. *Implicación institucional en Linked Open Data*, delimitado por los polos de 'Regulación' y 'Datos y Metadato')

- A.1. Políticas orientadas hacia datos y modelos de datos.
- A.2. Movimientos orientados hacia ampliar la apertura de datos.

B. *Oferta de Datos y Contenidos*, delimitado por 'Tecnologías, Plataformas' y 'Datos y Metadatos'.

- B.1. Esquemas de metadatos y vocabularios de valores publicados como datos enlazados.
- B.2. Datasets publicados de forma dinámica.

C. *Oferta de Servicios y Aplicaciones*, delimitado por los 'Tecnologías, Plataformas' y 'Aplicaciones de consumo'.

- C.1. Tecnologías de publicación.
- C.2. Tecnologías de agregación y consumo de datos.

D. *Mercado eficiente y seguro*, delimitado por 'Aplicaciones de consumo' y 'Regulación'.

- D.1. Aplicaciones y servicios libres.
- D.2. Aplicaciones y servicios comerciales.

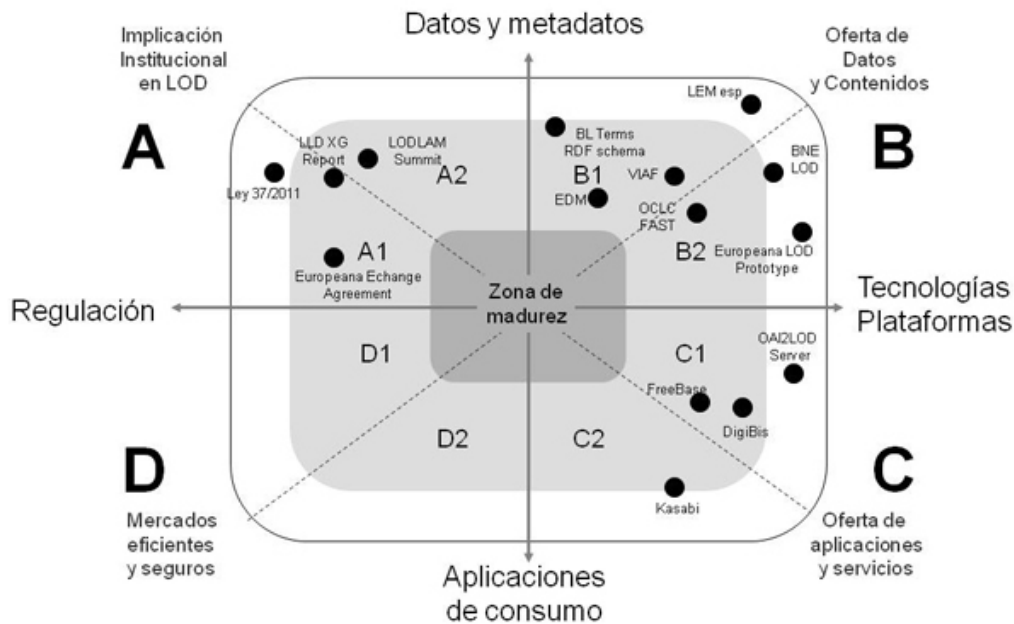


Figura 5. Aplicación del modelo en proyectos *linked open data*-libraries, archives and museums. Fuente: elaboración propia

En el esquema hemos situado, a modo de ejemplo, varios ejemplos del sector libraries, archives and museums, como VIAF (Virtual International Authority File), EDM (Europeana Data Model), o LEM (Lista de Encabezamientos de Materias para Bibliotecas Públicas), Digibis (Software para Biblioteca Digitales). Se observa que el cuadrante D (Mercados eficientes y seguros) es en el que menos ejemplos significativos se pueden encontrar aún: faltan aplicaciones, suficientemente difundidas y valoradas, que incorporen datos enlazados bibliotecarios.

La cercanía a los ejes significa que se observa mayor compenetración entre dos características próximas, y la proximidad al centro, una mayor alineación con el resto de los elementos, y por lo tanto, mayor relevancia en el ecosistema.

Tomemos como ejemplo un conjunto de datos publicado por una Biblioteca Nacional:

- El elemento se situaría inicialmente en el eje de Datos y Metadatos, equidistante de los ejes contiguos y alejado del punto de cruce.
- Si se basa en un marco jurídico claro que regula la difusión y el tipo de licencia, se representaría desplazado al cuadrante A.
- Si, por el contrario, el aspecto principal es la tecnología estandarizada y las mejores prácticas de publicación, y están bien integradas en los sistemas corporativos de manejo de datos primarios, el elemento se desplazaría hacia el cuadrante B.
- Si al mismo tiempo existen ya aplicaciones que integran sus datos en servicios al usuario final, se iría aproximando hacia el cuadrante D.
- Si las cuatro facetas estuvieran equilibradas, la ubicación del elemento tendería hacia la zona central, indicando mayor madurez, al producirse sinergias entre los diferentes factores de influencia.

Si, por ejemplo, pusieramos el foco en Europeana, podríamos preguntarnos de forma estructurada por los factores normativos y de financiación puestos en marcha por la Comisión Europea, la problemática con las licencias de reutilización, el volumen y calidad de datos, el acceso al contenido final, la adecuación de los repositorios nacionales al EDM (European Data Model), la existencia de SPARQL (Protocol and RDF Query Language) endpoints para su integración en otros webs, y la aparición de servicios o aplicaciones que consuman estos datos proporcionando utilidad al público en general e, incluso, modelos de negocio. También nos permitiría representar de forma diferenciada las facetas de Europeana como dataset y modelo de datos, como política de la agenda digital europea, como tecnologías de agregación o como aplicación de consumo. La existencia de elementos en cada uno de los cuadrantes, sería un indicador de una fase de mayor madurez y penetración de Europeana.

El modelo es una simplificación que busca facilitar la comprensión y comunicación, y no ha de tomarse como sustituto de otros instrumentos de análisis, como los catálogos e informes técnicos.

## Conclusiones.

El conjunto de todos los elementos expuestos hasta ahora podría inclinarnos a señalar la madurez de las tecnologías de datos enlazados, dado que en un conjunto de frentes (económicos, organizativos, técnicos, sociales y culturales) se están produciendo acciones sinérgicas a favor de

la mejora del acceso abierto a datos enlazados. Es arriesgado acertar a asignar el impacto a corto y medio plazo de los diferentes hitos que cierran el año 2011: OCLC hace accesible [FAST](#) (Faceted Application of Subject Terminology) como datos enlazados; la British Library libera la bibliografía nacional y sus [modelos de datos en RDE](#); [la Biblioteca Nacional de España](#) publica sus catálogos bibliográficos y de autoridades en RDF, utilizando las ontologías o vocabularios estándares de la IFLA (ISBD y familia derivada de los Functional Requirements); el Incubator Group sobre Datos Enlazados Bibliotecarios del W3c publica su informe final, la web de datos ocupa gran parte de las ponencias de [la Conferencia ELAG](#) (European Library Automation Group) bajo el lema '*It's the context, stupid*'; Europeana incorpora [linked open data](#) en su estrategia y desarrolla proyectos piloto, etc. Aunque resultan hitos significativos, vinculados a instituciones de alto nivel, denotan un movimiento de arriba abajo, faltaría todavía de la suficiente masa crítica y diversidad de actores para transformar el panorama real de los datos bibliotecarios en red. De hecho, el propio sector echa en falta un inventario de las aplicaciones de consumo, [aunque CKAN ha dado algún paso en este sentido](#), así como el gobierno británico en su portal '[Opening up government](#)'

No es fácil identificar el momento en que la dialéctica entre limitaciones y oportunidades se decanta a favor de la innovación. Las resistencias al cambio que presentan los sistemas socio-técnicos complejos a menudo reflejan unas utilidades que no se perciben conscientemente y que tienden a permanecer casi inalterables pese a las perspectivas optimistas, especialmente cuando el valor del cambio tiene que ser fuertemente percibido para que compense los altos costes de cualquier transición.

Observamos que los pasos hacia una fase de madurez de Linked Data en el sector libraries, archives and museums aún están demasiado escorados hacia la publicación de datos, en muchos casos en sus primeras versiones, y con muchos avances en la regulación y las estrategias de instituciones centrales de mayor entidad (Bibliotecas Nacionales y grandes consorcios). Se percibe aún la falta de tecnologías de referencia ampliamente aceptadas y de aplicaciones de consumo de datos bibliotecarios auténticamente demostrativas del valor que se le supone a los datos enlazados. Los sectores A y B de nuestro modelo están más desarrollados que los C y D. Nos inclinamos a considerar 2012 como un año clave para valorar si un crecimiento equilibrado entre ellos nos permite afirmar, en un plazo corto, que entramos en la fase de madurez para 2013.

No está de más recordar que los datos pueden jugar el papel de una nueva materia prima. Igual que con cada material se inicia un proceso de innovación para aprovechar sus cualidades, ¿está ocurriendo lo mismo con los datos? Ya no será sólo una cuestión de tecnologías, sino que entrará de lleno la administración de su propiedad y las restricciones sociales de uso. La abundancia de datos normalizados como factor de producción ayudará a definir nuevos espacios económicos y de valor. El ciclo de vida de los Linked Data, tanto si nos circunscribimos a una organización, como si hablamos de contextos sociales más amplios, implica un ciclo constante de re-equilibrio entre iniciativas para publicar datos, consumirlos, mejorar la semántica y calidad de los datos, definir políticas y marcos regulatorios y desarrollar tecnologías eficientes para gestionar la ingente cantidad de datos e interconexiones en sistemas distribuidos. Otros autores prefieren el término 'cadena de valor' ([Latif et al. 2009](#)) dando más importancia al progresivo enriquecimiento de los datos y funcionalidades mediante la participación de diferentes agentes hasta llegar a los usuarios finales.

La característica principal que cabe resaltar es que *linked open data* es un movimiento masivo y no sectorial, cuyos estándares de infraestructura pueden ser adoptados en campos dispares que van desde la información meteorológica a la catalogación bibliográfica. Por lo tanto las instituciones documentales debemos posicionar bien nuestras iniciativas, para que nuestros esfuerzos de publicación sintonicen con las tecnologías, las iniciativas de consumo y la regulación. Esto implica ampliar el abanico de acuerdos de colaboración más allá de las redes tradicionales, captando la atención de servicios afines con los que hasta ahora quizá no haya habido espacios de colaboración estrecha. Se adivina pues un período de comunicación de nuestras fuentes de información para conseguir que sean enlazadas. La construcción de un marco regulatorio a todos los niveles que sea propicio al uso de los datos puede estimular el surgimiento de servicios que generen valor usando nuestros datos. En este proceso se verán involucrados los agentes bibliotecarios tradicionales, pero también nuevos y diversos actores que pueden activarse al abrirse los datos de forma radical.

Los modelos revisados nos muestran la realidad económico-social de un fenómeno de red que va más allá de la simple producción de datos, entorno en que nuestro sector se inscribe. El análisis del escenario aumentado – sumando tecnologías, consumo y regulación– da cuenta de que estamos en los momentos iniciales de un fenómeno al que Glibraries, archives and museums todavía puede sumarse, pese a que la competencia y alternativas entre proveedores de datos aumentan.

Sería un ejercicio ilustrativo trazar un paralelismo con la extensión que el movimiento Open Access ha tenido en la última década. De la febril actividad de este campo pueden extraerse lecciones de interés para comprender los elementos clave para la consolidación de una tendencia tecnológica relacionada con la información. Sin que se trate de casos iguales, encontramos ciertos paralelismos entre este ejemplo y el despegue incipiente de *linked data* en nuestro sector.

Es importante profundizar nuestra comprensión de las dinámicas del auténtico ecosistema del espacio global de datos. La mera planificación no abarca todas las interacciones presentes, aunque las tendencias son el resultado de delicados equilibrios de éxito, adopción, experimentación y visibilidad. La evolución de linked data, y por consiguiente de la web semántica, es un asunto colaborativo, o si se quiere competitivo, en el que el encadenamiento de efectos positivos mutuos entre agentes es el que produce la masa crítica. Por ello no puede ser dirigida ni planificada con los instrumentos tradicionales de las políticas públicas, demasiado orientados a pautar los procesos.

*linked open data* va actuar de espoleta para la realización práctica de la web semántica, porque apoyada en los servicios e inversiones desarrollados con estas tecnologías, la web semántica se va a tangibilizar. La etiqueta 'Linked Open Data' puede, durante un periodo, 'oscurecer' a la web semántica, aunque esté asentando la infraestructura de base para construirla.

## Sobre los autores

**Tomás Saorín** es doctor en Documentación, y trabaja como documentalista en la Comunidad Autónoma de la Región de Murcia. Ha sido profesor asociado de la Facultad de Comunicación y Documentación de la Universidad de Murcia. Su email de contacto es [tsa@um.es](mailto:tsa@um.es).

**Fernanda Peset** es doctora en Documentación y profesora titular de la Universidad Politécnica de Valencia. Participa en proyectos como IraLIS, E-LIS y el Grupo Ciepi. Su email de contacto es: [mpesetm@upv.es](mailto:mpesetm@upv.es).

**Antonia Ferrer** es doctora en técnicas y métodos actuales en información y documentación, profesora titular de la Universidad Politécnica de Valencia (UPV) y coordinadora de investigación en Florida Universitaria. Ha coordinado numerosos proyectos nacionales e internacionales. Su email de contacto es: [anfesa@upv.es](mailto:anfesa@upv.es)

## Referencias

- Arjonilla Domínguez, S.J. y Medina Garrido, J.A. (2009). La gestión de los sistemas de información en la empresa: teoría y casos prácticos. Madrid: Pirámide.
- Abella, A. (2011). [Reutilización de información pública y privada en España: avance de situación para agentes públicos y privados. Una oportunidad para los negocios y el empleo](#). Madrid: Rooter Analysis. Mayo 2011, <http://www.rooter.es/node/87> (Archived by WebCite® at <http://www.webcitation.org/6CdebbuAN>)
- Allemang, D. y Hendler, J.A. (2008). Semantic Web for the working ontologist: effective modeling in RDFS and OWL. San Francisco, CA: Morgan Kaufmann; Oxford: Elsevier Science.
- Berners-Lee, T. (2006). [Linked data. World Wide Web Consortium design issues](#). <http://www.w3.org/DesignIssues/LinkedData.html> (Archived by WebCite® at <http://www.webcitation.org/6CdekvSsi>)
- Berners-Lee, T. (2009). Tim Berners-Lee on the next Web. [Video podcast]. Consultado el 08 de marzo 2013 de [http://www.ted.com/talks/lang/en/tim\\_berniers\\_lee\\_on\\_the\\_next\\_web.html](http://www.ted.com/talks/lang/en/tim_berniers_lee_on_the_next_web.html) (Archived by WebCite® at )
- Berners-Lee, T. (2010). [The year open data went worldwide](#). [Video podcast]. Consultado el 08 de marzo 2013 de [http://www.ted.com/talks/tim\\_berniers\\_lee\\_the\\_year\\_open\\_data\\_went\\_worldwide.html](http://www.ted.com/talks/tim_berniers_lee_the_year_open_data_went_worldwide.html) (Archived by WebCite® at <http://www.webcitation.org/6EyAa0ri6>)
- Berners-Lee, T. & Fischetti, M. (1999). *Weaving the Web: the original design and ultimate destiny of the World Wide Web*. New York, NY: Harper-Collins.
- Berners-Lee, T., Hendler, J. & Lassila, O. (2001). [The semantic Web](#). *Scientific American*, **284**(5), 34-56. Consultado el 08 de marzo 2013 de from <http://www.sciam.com/article.cfm?id=the-semantic-web> (Archived by WebCite® at <http://www.webcitation.org/6EyQVevVv>)
- Bizer, C., Cyganiak, R. y Heath, T. (2007). [How to publish linked data on the web](#). Retrieved 8 March, 2013 from <http://www4.wiwiw.fu-berlin.de/bizer/pub/LinkedDataTutorial/> (Archived by WebCite® at <http://www.webcitation.org/6CdeoBlkz>)
- Bizer, C., Heath, T. y Berners-Lee, T. (2009). Linked data, the story so far. *International Journal on Semantic Web and Information Systems*, **5**(3), 1-22
- Castells, M. (1996). *La era de la información: economía, sociedad y cultura*. Vol. 1. *La sociedad red*. Madrid: Alianza.
- Cornellá, A. y Flores, A. (2007). *La alquimia de la innovación*. Barcelona: Deusto Ediciones.
- Discovery Programme. (2012). [Discovery open metadata principles](#). Consultado el 08 de marzo 2013 de <http://discovery.ac.uk/businesscase/principles/> (Archived by WebCite® at <http://www.webcitation.org/6EyWZFYzU>)
- Gillespie, A. (2007). [Foundations of economics. Additional chapter on business strategy](#). Oxford University Press. Retrieved 8 March, 2013 from [http://fdslive.oup.com/www.oup.com/orc/resources/busecon/economics/gillespie\\_econ2e/01student/additional/business\\_strategy.pdf](http://fdslive.oup.com/www.oup.com/orc/resources/busecon/economics/gillespie_econ2e/01student/additional/business_strategy.pdf) (Archived by WebCite® at <http://www.webcitation.org/6F2rSYZon>)
- López Sánchez, J.I. y Arroyo Barrigüete, J.L. (2006). Externalidades de red en la economía digital: una revisión teórica. *Economía Industrial*, No. 361, 21-31
- Fenn, J. (2011). *Gartner's hype cycle special report for 2011*. Stamford, CT: Gartner.
- Fenn, J. y Raskino, M. (2008). *Mastering the hype cycle: how to choose the right innovation at the right time*. Boston, MA: Harvard Business Press.
- Hausenblas, M. (2009). [Linked data applications: the genesis and the challenges of using linked data on the web](#). Galway, Ireland: Digital Enterprise Research Institute. (DERI Technical Report 2009-07-26). Retrieved 8 March, 2013 from <http://linkeddata.deri.ie/tr/2009-ld2webapp> (Archived by WebCite® at <http://www.webcitation.org/6F2E35aXP>)
- Heath, T. y Bizer, C. (2011). [Linked data: evolving the web into a global data space](#). San Raphael, CA: Morgan & Claypool



- Publishers. (Synthesis Lectures on the Semantic Web: Theory and Technology, 2011, v.1, n. 1). Retrieved 8 March, 2013 from <http://linkeddatabook.com/book> (Archived by WebCite® at <http://www.webcitation.org/6CdeWUO6>)
- Hendler, J. y Golbeck, J. (2008). [Metcalf's law, Web 2.0, and the semantic web](#). *Journal of Web Semantics*, 6(1), 14-20. Retrieved 11 marh 2013 from <http://www.cs.umd.edu/~golbeck/downloads/Web20-SW-JWS-webVersion.pdf> (Archived by WebCite® at <http://www.webcitation.org/5bD4j1v3T>)
  - Herman, I. (2011). [Semantic Web adoption and applications](#). Consultado el 08 de marzo 2013 de <http://www.w3.org/People/Ivan/CorePresentations/Applications/Applications.pdf>
  - Latif, A., Saeed, A.U., Hoefler, P., Stocker, A. y Wagner, C. (2009). [The linked data value chain: a lightweight model for business engineers](#). In *Proceedings of the 5th International Conference on Semantic Systems, 2-4 September, Graz, Austria*. (pp. 568-575). Retrieved 11 march 2013 from [http://know-center.tugraz.at/download\\_extern/papers/ldvc.pdf](http://know-center.tugraz.at/download_extern/papers/ldvc.pdf) (Archived by WebCite® at <http://www.webcitation.org/6F2s0kMYZ>)
  - Moore, G.A. (2001). *Crossing the chasm: marketing and selling high-tech products to mainstream customers*. (Revised ed.). New York, NY: PerfectBound.
  - O'Reilly, T. (2005). [What is Web 2.0: design patterns and business models for the next generation of software](#). Sebastopol, CA: O'Reilly Media. Retrieved 8 March, 2013 from <http://oreilly.com/web2/archive/what-is-web-20.html> (Archived by WebCite® at <http://www.webcitation.org/6Cd8Xtpx>)
  - Pastor Sánchez, J.A. (2011). *Tecnologías de la web semántica*. Barcelona: Editorial Universitat Oberta de Catalunya.
  - Pedraza-Jiménez, R., Codina, L. y Rovira, C. (2007). [Web semántica y ontologías en el procesamiento de la información documental](#). *El profesional de la información*, 16(6), 569-578. Retrieved 11 March 2013 from <http://eprints.rclis.org/14298/1/webSemanticaOntologias2007.pdf> (Archived by WebCite® at <http://www.webcitation.org/6F2tj1XND>)
  - Peset, F., Ferrer-Sapena, A. y Subirats-Coll, I. (2011). [Open data y linked open data: su impacto en el área de bibliotecas y documentación](#). *El profesional de la información*, 10(2), 165-173. (Archived by WebCite® at <http://www.webcitation.org/6Cd8GsbLB>)
  - Phifer, G. (2011). *Hype cycle for web and user interaction technologies, 2011*. Stamford, CT: Gartner.
  - Pollock, R. (2011). [Building the \(open\) data ecosystem](#). [Web log post] Retrieved 8 March, 2013 from <http://blog.okfn.org/2011/03/31/building-the-open-data-ecosystem/> (Archived by WebCite® at <http://www.webcitation.org/6CefTj7HU>)
  - Poynder, R. (2011). [Suber: leader of a leaderless revolution](#). *Information Today*, 28(7). Retrieved 8 March, 2013 from <http://www.infoday.com/it/jul11/Suber-Leader-of-a-Leaderless-Revolution.shtml> (Archived by WebCite® at <http://www.webcitation.org/6CefRRXYE>)
  - Rogers, E.M. (2003). *Diffusion of innovations*. (5ª ed.). New York, NY: Free Press
  - Saorín, T. (2011). [Cómo linked open data impactará en las bibliotecas a través de la innovación abierta](#). *Anuario ThinkEPI*, 6, 288-292. Retrieved 8 March, 2013 from <http://www.thinkepi.net/como-linked-open-data-impactara-en-las-bibliotecas-a-traves-de-la-innovacion-abierta> (Archived by WebCite® at <http://www.webcitation.org/6CefOMeMD>)
  - Shapiro, C. y Varian, H.R. (2000). *El dominio de la información: Guía estratégica para la economía de la red*. Barcelona: Antoni Bosch Editor.
  - Sierra, D. (2011, May 11). [ImprimirAbredatos 2011, cómo poner información pública al servicio del ciudadano](#). [Web log post]. Retrieved 12 March, 2013 from <http://www.rtve.es/noticias/20110511/abredatos-2011-como-poner-datos-publicos-servicio-del-ciudadano/431358.shtml>
  - Smulders, C. (2011). [Magic quadrants: positioning technology players within a specific market](#). Stamford, CT: Gartner. Retrieved 8 March, 2013 from [http://www.gartner.com/technology/research/methodologies/research\\_mq.jsp](http://www.gartner.com/technology/research/methodologies/research_mq.jsp) (Archived by WebCite® at <http://www.webcitation.org/6CefL2J6R>)
  - Verdonckt, L. (2011, October 3). [First release of the LOD2 stack](#). [Web log post]. Consultado el 08 de marzo 2013 de <http://lod2.eu/BlogPost/677-first-release-of-the-lod2-stack.html> (Archived by WebCite® at <http://www.webcitation.org/6EyBNWdTW>)
  - W3C. *Library Linked Data Incubator Group*. (2011). [Final report](#). Consultado el 08 de marzo 2013 de <http://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/> (Archived by WebCite® at <http://www.webcitation.org/6EyVvQXgR>)
  - Wood, D. (ed.) (2010). [Linking enterprise data](#). 3RoundStones.com Retrieved 8 March, 2013 from [http://3roundstones.com/led\\_book/led-contents.html](http://3roundstones.com/led_book/led-contents.html) (Archived by WebCite® at <http://www.webcitation.org/6CefldLXA>)

### Abstract in english

**Introduction.** The paper develops a framework for identifying the level of penetration and maturity of *linked data* technologies, as a specific facet of the W3C open model for the semantic Web, in the field of libraries, archives and museums.

**Method.** The paper applies the conceptual framework of the classical theory of diffusion of innovations, along with patterns of life-cycle analysis of technologies (hype cycle) developed by Gartner Group. We compose a model for the analysis of global factors, adapted to the nature of *linked data* technology.

**Results.** A framework for identifying the level of maturity of the linked data technology drivers is outlined. Our analysis indicates that, despite the intense activity detected in the libraries and archives sector, there is still not enough balance between publication, consumption, regulation and technologies to state that we are entering in the phase of maturity.

**Conclusions.** More linked data initiatives aimed at apps and consumer data are needed, besides exploiting collaborative aspects and build further networks with entities outside the field of libraries, archives and museums. We assume that leadership from the biggest institutions will continue and open innovation will be intensified.

#### [How to cite this paper](#)

Saorín, T., Peset, F. & Ferrer-Sapena, A. (2013). Factores para la adopción de *linked data* e implantación de la web semántica en bibliotecas, archivos y museos. *Information Research*, **18**(1) paper 570. [Available at <http://InformationR.net/ir/18-1/paper570.html>]

#### Find other papers on this subject

[Scholar Search](#)[Google Search](#)[Bing](#)

Check for citations, [using Google Scholar](#)

Like { 4 } [Twitter](#) { 0 }

**1307**

© the authors, 2013.

Last updated: 8 March, 2013

- [Contents](#) |
- [Author index](#) |
- [Subject index](#) |
- [Search](#) |
- [Home](#)